

**Manuscript version: Working paper (or pre-print)**

The version presented here is a Working Paper (or 'pre-print') that may be later published elsewhere.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/136879>

**How to cite:**

Please refer to the repository item page, detailed above, for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# JNCD-BASED PERCEPTUAL COMPRESSION OF RGB 4:4:4 IMAGE DATA

*Lee Prangnell and Victor Sanchez*  
{lee.prangnell, v.f.sanchez-silva}@warwick.ac.uk

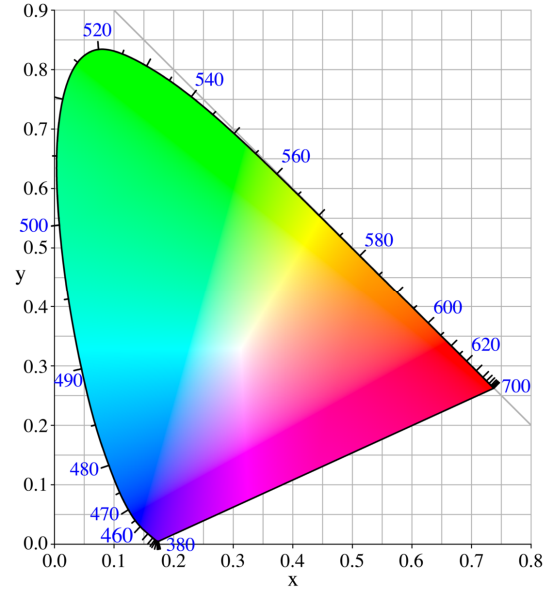
Department of Computer Science, University of Warwick, England, UK

## ABSTRACT

In contemporary lossy image coding applications, a desired aim is to decrease, as much as possible, bits per pixel without inducing perceptually conspicuous distortions in RGB image data. In this paper, we propose a novel color-based perceptual compression technique, named RGB-PAQ. RGB-PAQ is based on CIELAB Just Noticeable Color Difference (JNCD) and Human Visual System (HVS) spectral sensitivity. We utilize CIELAB JNCD and HVS spectral sensitivity modeling to separately adjust quantization levels at the Coding Block (CB) level. In essence, our method is designed to capitalize on the inability of the HVS to perceptually differentiate photons in very similar wavelength bands. In terms of application, the proposed technique can be used with RGB (4:4:4) image data of various bit depths and spatial resolutions including, for example, true color and deep color images in HD and Ultra HD resolutions. In the evaluations, we compare RGB-PAQ with a set of anchor methods; namely, HEVC, JPEG, JPEG 2000 and Google WebP. Compared with HEVC HM RExt, RGB-PAQ achieves up to 77.8% bits reductions. The subjective evaluations confirm that the compression artifacts induced by RGB-PAQ proved to be either imperceptible (MOS = 5) or near-imperceptible (MOS = 4) in the vast majority of cases.

## 1. INTRODUCTION

Attaining visually lossless quality at a low bit size has always been a topic of interest in the image coding community [1]. This is indeed the case in chromatic still image coding including medical image compression and the coding of camera-captured (natural) image content. The viewing of compressed RGB data on ITU-R BT.709 High Definition (HD) and ITU-R BT.2020 Ultra HD Visual Display Units (VDUs) [2, 3] is becoming increasingly ubiquitous on a global scale. For example, 24-bit RGB 4:4:4 data — of various resolutions — is used in a multitude of fields including digital entertainment, computer vision, medical imaging [4] and tone mapping in High Dynamic Range (HDR) digital photography. The increasing prevalence of RGB-based multimedia playback on state-of-art VDUs makes perceptual compression increasingly desirable. This is primarily the case because image compression artifacts, such as blocking, blurring and ringing artifacts, are typically more noticeable on state-of-the-art HD 1080p and Ultra HD 2160p VDUs [5]–[7]. This is especially the case if, for example, a moderately to coarsely quantized 720p image is displayed on an Ultra HD VDU in full screen mode. Regarding the lossy coding of RGB data, image compression standards in their present standardized form are not perceptually optimized. For instance, popular compression platforms, including JPEG (ITU-T T.81) [8] and JPEG 2000 (ISO/IEC 15444-1:2019) [9], take neither JNCD nor HVS spectral sensitivity modeling into account. This provides an opportunity to fill a research gap in terms of potentially achieving visually lossless (or near visually lossless) quality in addition to attaining noteworthy bit savings.



**Fig 1:** CIE chromaticity diagram [10] that shows the full range of chroma, hue and saturation that the HVS is capable of perceiving. The values in blue refer to the wavelength of photons (in nm).

There are several lossy image compression platforms in existence. Some popular examples include the aforementioned JPEG [8] and JPEG 2000 [9] standards in addition to Google’s WebP offering [11]. A recent addition to these widely used image coding platforms is Better Portable Graphics (BPG) [12]. BPG is based on the HEVC standard (ITU-T H.265 [13]). In BPG, the HEVC-based compression algorithms, including intra coding, are integrated into BPG from the open source HEVC x265 codec. Almost three decades since its inception, JPEG, which, to reiterate, is not perceptually optimized, is still the de facto lossy image compression format utilized on the Internet. JPEG 2000 is mostly employed in lossy and lossless medical image applications, such as Whole Slide Image (WSI) coding. Google WebP is the default image format in Google Chrome-based environments; WebP is based on the VP8 video codec [11].

In this paper, we propose a novel JNCD-based perceptual quantization technique, named RGB-PAQ, which operates at the Coding Block (CB) level in HEVC HM RExt. In terms of the core functionality of RGB-PAQ, we employ JNCD-based color masking in addition to accounting for HVS spectral sensitivity (that is, quantizing data in the B and R channels more coarsely than data in the G channel). The primary justification for this is the fact that the HVS is poor at detecting small differences in color. This could be in the form of slightly different shades of the same color or different colors that look very similar (e.g., aqua versus turquoise); see Fig. 1. As compared with well-established and emerging image compression techniques, RGB-PAQ offers the key advantages of quality preservation in addition to noteworthy bit reductions.

The rest of this paper is organized as follows. Section 2 includes an overview of lossy and perceptual image coding. We provide detailed information on the proposed RGB-PAQ technique in Section 3. The evaluation, results and discussion of RGB-PAQ is shown in Section 4. Finally, section 5 concludes the paper.

## 2. LOSSY AND PERCEPTUAL IMAGE CODING

In terms of lossy coding, perceptual considerations have always been integral to the official development and standardization of image and video coding platforms. In JPEG [8], JPEG 2000 [9], WebP [11] and HEVC [13, 14], the key lossy compression methods are transform coding and scalar quantization; scalar quantization is the main lossy compression method in the image coding pipeline. The perceptual aspect of transform coding and quantization is related to the Modulation Transfer Function (MTF), which is based on the Contrast Sensitivity Function (CSF). In terms of MTF modeling, several experiments in the field of visual psychophysics have shown that the HVS is less sensitive to changes in high spatial frequency detail. In transform coding, the high spatial frequency detail in an image corresponds to high frequency transform coefficients. Therefore, in lossy image coding, quantization systems are designed primarily to target the high frequency coefficient sub-band. All modern image and video coding platforms, including JPEG, JPEG 2000 and HEVC, follow the MTF model intrinsic to transform coding and quantization.

Lossy image and video compression is, in essence, all about the way in which the HVS interprets physical photon waves and physical luminance displayed on VDUs. The HVS perceives the combination of photon waves and luminance as a vast range of colors; note that color is a subjective phenomenon and does not exist outside the perceptual domain. In photobiology, cone sensitivity experiments have revealed that 64% are sensitive to photons perceived as red, 32% perceived as green and 4% perceived as blue; this is known as trichromatic color vision. Although there are more cones that are sensitive to photons interpreted as red, the HVS is much more sensitive to the perceived brightness of photons that are interpreted as green (see Fig. 1). Therefore, the green channel in the RGB color space is the most perceptually important and thus has the largest weight when converting from RGB to YCbCr (i.e., the Y channel is correlated to the G channel).

Regarding the modeling of perceptual image compression techniques, the Weber-Fechner law ([15]) confirms that there is a mathematical relationship between the subjective sensation of a physical stimulus and the intensity of the actual physical stimulus. This implies that there is a mathematical relationship between the perception of brightness and the intensity of physical luminance in nature. Likewise, it also implies that there is a mathematical relationship between perceived color (i.e., chroma, hue, saturation and contrast) and photon waves in nature. Therefore, the scientific basis of the Weber-Fechner law is very useful for perceptual compression research. In the context of JNCD-based modeling and perceptual compression, the HVS is poor at distinguishing small differences in color. This fact has given rise to perceptually uniform color spaces including CIELAB [16]. To reiterate, the HVS is most sensitive to photon waves — and the associated luminance — that are perceived as green [17]. RGB-PAQ therefore treats the G channel as the most perceptually important channel during the CB-level perceptual quantization process. We utilize the CIELAB JNCD threshold value in order to determine the perceived acceptability of quantization-induced compression artifacts in a lossy-coded picture. In other words, we model perceptual image coding in RGB-PAQ by accounting for the way that the HVS interprets the structure of brightness, chroma, hue, saturation and contrast.

---

### Algorithm 1: Iterative Process for CB QP Increments (when $\Delta E_{ab} < 2.3$ )

---

```

1: procedure Perceptual_CB_QP( $Q_G, Q_B, Q_R$ )
2:   while  $\Delta E_{ab} < 2.3$  do
3:     CB-Level_QP_Incrementation:
4:       repeat
5:         Increment_Blue_CB-Level_QP: // increment blue CB QP first
6:          $i_B = 1; Q_B = (Q_B + i_B)$ 
7:         do  $i_B++$  until  $i_B = 6$ 
8:         if  $\Delta E_{ab} \approx 2.3$ , then goto End:
9:         else goto Increment_Red_CB-Level_QP:
10:        Increment_Red_CB-Level_QP: // increment red CB QP second
11:         $i_R = 1; Q_R = (Q_R + i_R)$ 
12:        do  $i_R++$  until  $i_R = 6$ 
13:        if  $\Delta E_{ab} \approx 2.3$ , then goto End:
14:        else goto Increment_Green_CB-Level_QP:
15:        Increment_Green_CB-Level_QP: // increment green CB QP last
16:         $i_G = 1; Q_G = (Q_G + i_G)$ 
17:        do  $i_G++$  until  $i_G = 3$ 
18:        if  $\Delta E_{ab} \approx 2.3$ , then goto End:
19:        else goto CB-Level_QP_Incrementation:
20:      until  $\Delta E_{ab} \approx 2.3$ 
21:    End:
22:  end while
23: end procedure

```

---

## 3. PROPOSED RGB-PAQ TECHNIQUE

RGB-PAQ is integrated into JCT-VC HEVC HM RExt [18]. RGB-PAQ operates at the CB level for performing perceptual quantization operations and at the Coding Unit (CU) level for computing JNCD color difference measurements. With relevance to RGB-PAQ, the direct coding of RGB 4:4:4 image and video data — of various bit depths — is a core feature in the Range Extensions of the HEVC standard (i.e., JCT-VC HEVC HM RExt). As is the case with the coding of YCbCr 4:4:4 image and video data in HEVC, the raw RGB 4:4:4 data is partitioned into RGB Coding Units (CUs). Each CU contains three equal sized CBs — i.e., a Red CB, a Green CB and a Blue CB — due to the absence of chroma subsampling [19]. In the lossy coding pipeline [20], the default scalar quantization techniques are known as Uniform Reconstruction Quantization (URQ) [21, 22] and Rate Distortion Optimized Quantization (RDOQ) [23]. Both URQ and RDOQ are not perceptually optimized. They do not account for the color related perceptual redundancies that are present in each R, G and B color channel (the B and R color channels in particular). The proposed RGB-PAQ technique solves this problem.

In the proposed method, we employ the CIELAB color difference formula  $\Delta E_{ab}$ , as shown in [24], to compute the JNCD threshold value.  $\Delta E_{ab}$  constitutes the Euclidean distance between two colors in the CIELAB color space [25]. Experiments in the field of colorimetry show that  $\Delta E_{ab} \approx 2.3$  equates to the JNCD threshold [26]. We utilize  $\Delta E_{ab}$  for computing the CIELAB color difference between raw pixels and reconstructed pixels at the CU level in the HEVC encoder loop. After this process, the levels of CB-level quantization are adjusted until  $\Delta E_{ab} \approx 2.3$ . In terms of computing  $\Delta E_{ab}$ , let  $C_w$  and  $C_r$  denote two sets, both of which contain three values.  $C_w$  and  $C_r$  denote the raw CU and the reconstructed CU, respectively, containing the mean values for each constituent R, G and B CB. Therefore,  $C_w = \{\mu G_w, \mu B_w, \mu R_w\}$  and  $C_r = \{\mu G_r, \mu B_r, \mu R_r\}$ . Variables  $\mu G_w, \mu B_w$  and  $\mu R_w$  refer to the mean sample values in each raw G CB, B CB and R CB. Conversely, variables  $\mu G_r, \mu B_r$  and  $\mu R_r$  refer to the mean sample values in each reconstructed G CB, B CB and R CB. The  $\Delta E_{ab}$  formula in [24] is employed to compare the aforementioned raw RGB CU  $C_w$  with the reconstructed RGB CU  $C_r$ . For example, if  $\Delta E_{ab} < 2.3$  when comparing  $C_w$  with  $C_r$ , we increment CB-level quantization until  $\Delta E_{ab} \approx 2.3$  (see Algorithm 1 and Fig. 2).

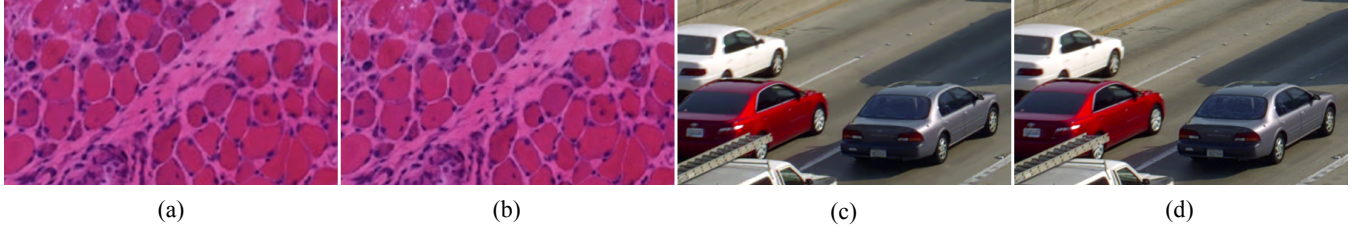
181	167	245	192		25	81	32	10		11	100	222	19
193	41	87	65		43	16	18	15		45	88	91	82
70	30	149	129		91	20	16	45		63	75	150	14
174	127	57	178		14	44	75	89		12	116	41	99

(a<sub>1</sub>)  $G_w$                       (a<sub>2</sub>)  $B_w$                       (a<sub>3</sub>)  $R_w$

170	180	255	210		15	60	25	20		4	115	200	35
160	32	115	40		60	30	10	24		30	80	100	72
52	20	125	110		80	30	30	32		74	86	162	10
200	100	70	160		31	66	57	105		18	116	50	85

(b<sub>1</sub>)  $G_r$                       (b<sub>2</sub>)  $B_r$                       (b<sub>3</sub>)  $R_r$

**Fig. 2:** Toy 4×4 raw G, B and R CBs in subfigures (a<sub>1</sub>), (a<sub>2</sub>) and (a<sub>3</sub>) within a raw CU ( $C_w$ ) and toy 4×4 reconstructed G, B and R CBs in subfigures (b<sub>1</sub>), (b<sub>2</sub>) and (b<sub>3</sub>) within a reconstructed CU ( $C_r$ ). Though CUs and the constituent CBs in HEVC are partitioned to the sample ranges of 8×8 to 64×64 [19], this figure of toy 4×4 CUs is shown to illustrate how we compute the JNCD threshold (i.e.,  $\Delta E_{ab} \approx 2.3$ ). That is, the mean sample values of the raw and reconstructed CBs, respectively, are derived from  $G_w$ ,  $B_w$ ,  $R_w$  — as shown in subfigures (a<sub>1</sub>), (a<sub>2</sub>) and (a<sub>3</sub>) — and  $G_r$ ,  $B_r$  and  $R_r$  — as shown in subfigures (b<sub>1</sub>), (b<sub>2</sub>) and (b<sub>3</sub>). Once the mean CB values have been computed for each raw and reconstructed CU ( $C_w$  and  $C_r$ , respectively),  $\Delta E_{ab}$  is then computed for each CU in the HEVC encoder loop until  $\Delta E_{ab} \approx 2.3$  is reached.



**Fig. 3:** Subfigure (a) RGB-PAQ coded 4K medical image (WSI) versus subfigure (b) HEVC-coded medical image (WSI). Compared with the raw data, both RGB-PAQ and HEVC achieve MOS = 5 (majority of tests) and SSIM  $\geq 0.99$ . Subfigure (c) RGB-PAQ coded natural content (Traffic) versus subfigure (d) JPEG-coded natural content (Traffic); compared with the raw data, both RGB-PAQ and JPEG achieve MOS = 5 (majority of tests) and SSIM  $\geq 0.99$ .

Regarding the way in which RGB-PAQ performs CB-level Quantization Parameter (QP) increments on 8×8 to 64×64 CBs, it is important to note that HVS spectral sensitivity, as explained previously, governs the order as to which CB within a CU is quantized first. In other words, if  $\Delta E_{ab} < 2.3$ , the proposed method firstly increments the B CB QP, then the R CB QP, then finally the G CB QP until  $\Delta E_{ab} \approx 2.3$  (see Algorithm 1 for the QP increment values). RGB-PAQ operates in this manner because, as previously stated, the HVS is most sensitive to photons perceived as green [27]; therefore, G CBs are quantized less aggressively. Note that if  $\Delta E_{ab} > 2.3$  due to a high initial QP setting, then RGB-PAQ subsequently decreases CB-level quantization. That is, instead of incrementing the CB-level QPs, they are instead decremented in the order of G, R, B (again, according to HVS spectral sensitivity). In other words, the QP for the G CB is decremented first, then the R CB QP and finally the B CB QP until  $\Delta E_{ab} \approx 2.3$ .

In the proposed RGB-PAQ algorithm, the CB-level perceptual QPs, denoted as  $Q_G$ ,  $Q_B$  and  $Q_R$ , and the corresponding CB-level Quantization Step Sizes (QSteps), denoted as  $S_G$ ,  $S_B$  and  $S_R$ , are shown in (1)–(6), respectively:

$$Q_G(S_G) = \left\lceil \left[ 6 \times \log_2(S_G) \right] + 4 \pm i_G \right\rceil \quad (1)$$

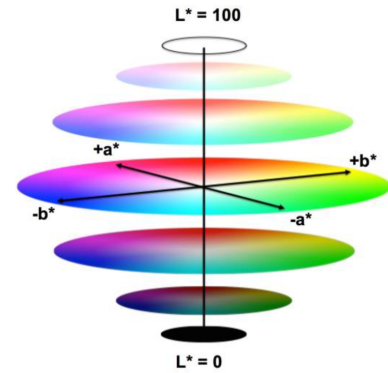
$$S_G(Q_G) = 2^{\frac{Q_G - 4 \pm i_G}{6}} \quad (2)$$

$$Q_B(S_B) = \left\lceil \left[ 6 \times \log_2(S_B) \right] + 4 \pm i_B \right\rceil \quad (3)$$

$$S_B(Q_B) = 2^{\frac{Q_B - 4 \pm i_B}{6}} \quad (4)$$

$$Q_R(S_R) = \left\lceil \left[ 6 \times \log_2(S_R) \right] + 4 \pm i_R \right\rceil \quad (5)$$

$$S_R(Q_R) = 2^{\frac{Q_R - 4 \pm i_R}{6}} \quad (6)$$



**Fig. 4:** Conceptual diagram of the CIELAB color space [28]. Note how CIELAB uses a 3D coordinate system, which separates the lightness of color from the chroma, hue and saturation of color.

where  $i_G$ ,  $i_B$  and  $i_R$  refer to the incremental (or decremental) values for increasing (or decreasing) the G CB, B CB and R CB QPs accordingly. If  $\Delta E_{ab} < 2.3$ , then  $i_G = 1$ ,  $i_B = 1$  and  $i_R = 1$ . Conversely, if  $\Delta E_{ab} > 2.3$ , then  $i_G = -1$ ,  $i_B = -1$  and  $i_R = -1$ . Algorithm 1 highlights the iterative process for incrementing the CB-level QP values by  $i_G$ ,  $i_B$  and  $i_R$  (i.e., when  $\Delta E_{ab} < 2.3$ ). The CIELAB  $\Delta E_{ab}$  formula, as shown in [24], is computed in (7):

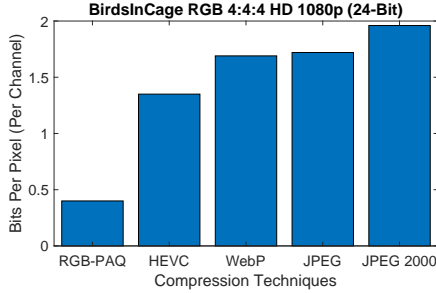
$$\Delta E_{ab} = \sqrt{(L_2 - L_1)^2 + (a_2 - a_1)^2 + (b_2 - b_1)^2} \quad (7)$$

where  $L$ ,  $a$  and  $b$  refer to Cartesian coordinates (see Fig. 4). In CIELAB,  $L$  constitutes the lightness of a color;  $L \in [0, 100]$  (value 0 represents black and value 100 represents white). Coordinates  $a$  and  $b$  refer to the hue and saturation of a color. The chromatic  $a$  axis extends from green to red; this is typically denoted as  $(-a)$  for green and  $(+a)$  for red. Likewise, the chromatic  $b$  axis extends from blue to yellow, which is denoted as  $(-b)$  for blue and  $(+b)$  for yellow (see Fig. 4). The specific computations for  $L$ ,  $a$  and  $b$  are shown in [24]. Note that CIELAB is a perceptually uniform color space that is RGB color space independent [28]; therefore, CIELAB and  $\Delta E_{ab}$  are very useful for perceptual image coding.



**Table 1:** In green text, the bits per pixel (per channel) achieved by RGB-PAQ and each anchor technique; namely, HEVC, JPEG, JPEG 2000 and WebP. In blue text, the subjective visual quality is quantified by the Mean Opinion Score (MOS) — rounded to the nearest integer — for RGB-PAQ and anchors.

Bits Per Pixel, Per Channel (BPP) for RGB-PAQ versus Anchors and MOS Scores for RGB-PAQ versus Anchors										
	BPP	BPP	BPP	BPP	BPP	Rounded Mean Opinion Score (MOS)				
RGB Data	RGB-PAQ	HEVC	JPEG	JPEG 2000	WebP	RGB-PAQ	HEVC	JPEG	JPEG 2000	WebP
BirdsInCage	0.40	1.35	1.72	1.96	1.69	5	5	5	5	5
Bubbles	0.51	1.30	2.30	1.70	1.64	5	5	5	5	5
CrowdRun	2.16	5.86	5.38	6.93	6.53	4	5	5	4	5
CT	0.32	0.50	1.12	0.69	0.53	5	5	5	5	5
DuckAndLegs	2.21	5.20	5.75	6.91	5.27	5	5	5	5	5
Kimono	0.50	2.26	2.26	2.26	2.40	4	5	5	4	5
OldTownCross	1.30	5.24	4.49	6.06	5.89	4	5	5	4	5
ParkScene	1.10	4.01	3.53	4.05	3.58	4	5	5	4	5
Seeking	1.71	5.78	6.36	5.20	6.32	5	5	5	5	5
Traffic	1.14	2.18	3.06	3.69	2.84	5	5	5	5	5
VenueVu	0.64	1.12	1.53	2.22	1.39	4	5	5	4	5
WSI (4K)	0.53	0.75	0.94	1.21	0.95	5	5	5	5	5



**Fig. 5:** A bar graph showing the bit reductions (in bits per pixel, per channel) achieved by RGB-PAQ in comparison with HEVC, WebP, JPEG and JPEG 2000 on the BirdsInCage RGB 4:4:4 1080p 24-bit image.

#### 4. EVALUATION, RESULTS AND DISCUSSION

We implement the proposed RGB-PAQ technique into the JCT-VC HEVC HM 16.7 RExt codec [29]. RGB-PAQ is compared with well-established compression methods (anchors). The anchors include HEVC HM 16.7 RExt as well as the latest versions of the following open source codecs implemented in the FFMPEG 4 software [30]: Google WebP (libwebp), JPEG (mjpeg) and JPEG 2000 (jpeg2000). RGB-PAQ and anchors are tested on 12 raw RGB 4:4:4 images, all of which are HD 1080p except for CT (2000×2000) and WSI (Ultra HD 2160p). RGB-PAQ achieves a reconstruction quality of  $SSIM \geq 0.99$  in all tests. This score is quantified by the SSIM metric [31] that is included in the Video Multimethod Assessment Fusion (VMAF) [32] library in FFMPEG 4. To ensure that fair testing is achieved, we compressed all test images using the anchor methods to a reconstruction quality of  $SSIM \geq 0.99$ . Regarding the SSIM metric, a score of  $SSIM \geq 0.99$  is often considered to correlate with  $MOS \approx 5$ . Note that  $SSIM = 1$  is the maximum SSIM value (i.e., equates to lossless compression).

The subjective assessment — quantified using the Mean Opinion Score (MOS) as per ITU-T Rec. P.910 [33] — is arguably the most important aspect of the evaluation. Note that  $MOS = 5$  equates to visually lossless quality and  $MOS = 4$  corresponds to very slight perceptible distortion [33]. ITU-T Rec. P.910 advises a minimum of four participants. Therefore, five participants engaged in the subjective evaluations on HD VDUs at a viewing distance of approximately 0.75m in accordance with ITU-T Rec. P.910 recommendations. The participants carried out a total of 168 visual quality assessments. Over 58% of the participants reported no visual quality differences between RGB-PAQ coded sequences and the raw RGB data (i.e.,  $MOS = 5$ ); see Table 1 and also Fig. 6 for a visual example. The subjective evaluations provide evidence that RGB-PAQ successfully achieves visually lossless (or near visually lossless) quality in all tests conducted.



**Fig. 6:** A section from the BirdsInCage RGB 4:4:4 1080p image. Subfigure (a) shows the RGB-PAQ coded version and subfigure (b) shows the version from the raw RGB data. Subfigure (b) is 8 bits per pixel (per channel) and the RGB-PAQ coded version in (a) is 0.4 bits per pixel (per channel).

In comparison with anchors (see Table 1), RGB-PAQ attains vast reductions in terms of bits per pixel (per channel). The most noteworthy bit savings are accomplished on the BirdsInCage RGB 4:4:4 HD 1080p image (see Fig. 5 and Fig. 6). Compared with the HEVC HM RExt reference software [29], RGB-PAQ attains 0.40 bits per pixel (per channel), whereas HEVC achieves 1.35 bits per pixel (per channel). The most impressive compression results are obtained on images with a larger number of low variance CUs including BirdsInCage, Bubbles, CT, VenueVu and WSI. This is because the low variance CBs (the B and R CBs, in particular) are quantized more aggressively according to JNCD and HVS spectral sensitivity. The more coarsely quantized B and R channels are, in essence, perceptually masked by the less aggressively quantized G channel; this constitutes color masking. Finally, we achieved the impressive bits per pixel savings because of the JNCD and HVS spectral sensitivity modeling used in RGB-PAQ. In other words, by utilizing the CIELAB JNCD threshold (i.e.,  $\Delta E_{ab} \approx 2.3$ ) in addition to perceptually adjusting the QP at the CB level (in a specific order), we have shown that visually lossless compression can be achieved at a very low bits per pixel cost.

#### 5. CONCLUSION

In this paper, we propose a perceptual compression technique, named RGB-PAQ, for application with RGB 4:4:4 image data. In the proposed method, we exploit HVS spectral sensitivity and JNCD-based modeling in order to guide quantization adjustments at the CB level. Compared with anchors, RGB-PAQ attains vast bit savings, of up to 79.9%, while also achieving visually lossless quality (i.e.,  $MOS = 5$  and  $SSIM \geq 0.99$ ) in over 58% of subjective test cases. Regarding encoding and decoding time performances, no significant differences were recorded between RGB-PAQ and the anchor techniques.

## REFERENCES

- [1] H. R. Wu, A. R. Reibman, W. Lin, F. Pereira, and S. S. Hemami, "Perceptual Visual Signal Compression and Transmission," *Proc. IEEE*, vol. 101, no. 9, pp. 2025-2043, 2013.
- [2] ITU-R: Rec. BT.2020-2 (10/2015), "Parameter values for ultra-high definition television systems for production and international programme exchange," *ITU-R*, 2015.
- [3] ITU-R: Rec. 2100-2 (07/2018), "Image parameter values for high dynamic range television for use in production and international programme exchange," *ITU-R*, 2018.
- [4] V. Sanchez and M. Hernández-Cabronero, "Graph-Based Rate Control in Pathology Imaging With Lossless Region of Interest Coding," *IEEE Trans. Medical Imaging*, vol. 37, no. 10, pp. 2211-2223, 2018.
- [5] N. Casali, M. Naccari, M. Mrak and R. Leonardi, "Adaptive Quantisation in HEVC for Contouring Artefacts Removal in UHD Content," *IEEE Int. Conf. Image Processing*, 2015, pp. 2577-2581.
- [6] S. Daly and X. Feng, "Bit-depth extension using spatiotemporal microdither based on models of the equivalent input noise of the visual system," *Proc. SPIE*, 2003, vol. 5008, pp. 455-466.
- [7] T. K. Tan and Y. Suzuki, "Contouring artefact and solution," in *JCTVC-K0139, JCT-VC*, Shanghai, China, 2012.
- [8] ITU, "Information Technology – Digital Compression And Coding of Continuous-Tone Still Images – Requirements and Guidelines (JPEG)," T.81, *ITU-T*, 1992.
- [9] ISO/IEC 15444-1:2019, "Information technology — JPEG 2000 image coding system — Part 1: Core coding system," ISO/IEC, 2019.
- [10] Wikimedia, CIE Chromaticity Image [Online]. Available: <https://commons.wikimedia.org/wiki/File:CIExy1931.png>
- [11] J. Bankoski, P. Wilkins, and Y. Xu, "Technical overview of VP8, an open source video codec for the web," *IEEE Int. Conf. on Multimedia and Expo*, 2011, pp. 1-6.
- [12] F. Bellard, The BPG Image Format [Online]. Available: [bellard.org/bpg/](http://bellard.org/bpg/)
- [13] ITU-T: Rec. H.265/HEVC (Version 5) | ISO/IEC 23008-2, Information Technology – Coding of Audio-visual Objects, *JCT-VC (ITU-T/ISO/IEC)*, 2018.
- [14] G. Sullivan, J.-R. Ohm, W. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649-1668, 2012.
- [15] G. T. Fechner, "Elements of Psychophysics, Volume 1," (Translated by H.E. Adler), New York: Holt, Rinehart & Winston, 1860.
- [16] Specification ICC.1:2004-10, "Image technology colour management — Architecture, profile format, and data structure, *International Color Consortium*, 2004.
- [17] K. R. Gegenfurtner, "Cortical Mechanisms of Colour Vision," *Nature Neuroscience*, vol. 4, pp. 563-572, 2003.
- [18] D. Flynn, D. Marpe, M. Naccari, T. Nguyen, C. Rosewarne, K. Sharman, J. Sole and J. Xu, "Overview of the Range Extensions for the HEVC Standard: Tools, Profiles, and Performance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 4-19, 2016.
- [19] I.-K. Kim, J. Min, T. Lee, W.-J. Han and J. Park, "Block Partitioning Structure in the HEVC Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1697-1706, 2012.
- [20] M. Wein, "Residual Coding," in *High Efficiency Video Coding: Coding Tools and Specification*, Springer, 2015, pp. 205-227.
- [21] M. Budagavi, A. Fuldseth, G. Bjøntegaard, V. Sze and M. Sadafale, "Core Transform Design in the High Efficiency Video Coding (HEVC) Standard," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1649-1668, 2013.
- [22] M. Wein, "Quantizer Design," in *High Efficiency Video Coding: Coding Tools and Specification*, Springer, 2015, pp. 213-214.
- [23] M. Karczewicz, Y. Ye and I. Chong, "Rate Distortion Optimized Quantization," in *VCEG-AH21 (ITU-T SG16/Q6 VCEG)*, Antalya, Turkey, 2008.
- [24] M. R. Luo, G. Cui and B. Rigg, "The development of the CIE 2000 colour-difference formula: CIEDE2000," *Color Research and Application*, vol. 26, no. 5, pp. 340-350, 2001.
- [25] A. R. Robertson, "Historical development of CIE recommended color difference equations," *Color Research and Application*, vol. 15, no. 3, pp. 167-170, 1990.
- [26] G. Sharma, R. Bala, "Color Fundamentals for Digital Imaging," in *Digital Color Imaging Handbook*, CRC Press, 2002, pp. 31.
- [27] A. Chaparro, C. F. Stromeier, E. P. Huang, R. E. Kronauer and R. T. Eskew, Jr., "Colour is what the eye sees best," *Nature*, vol. 361, pp. 348-350, 1993.
- [28] Sappi Limited, "Defining and Communicating Color: The CIELAB System," Technical Document, pp. 5, 2013.
- [29] Joint Collaborative Team on Video Coding (JCT-VC). JCT-VC HEVC HM Reference Software, HM 16.7 [Online]. Available: [hevc.hhi.fraunhofer.de/](http://hevc.hhi.fraunhofer.de/)
- [30] FFMPEG Team, FFMPEG Multimedia Framework v4 [Online]. Available: [github.com/FFmpeg/FFmpeg/](https://github.com/FFmpeg/FFmpeg/)
- [31] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [32] R. Rassool, "VMAF reproducibility: Validating a perceptual practical video quality metric," *IEEE Int. Symp. on Broadband Multimedia Systems and Broadcasting*, 2017, pp. 1-2.
- [33] ITU-T, "Rec. P.910: Subjective video quality assessment methods for multimedia applications," *ITU-T*, 2008.